

CLAIMS

1. A server/client system in which a plurality of servers and a plurality of clients are connected through a network, and the servers execute a process based on a process request from the client and transmit a process result to the client, wherein
- at least one of the servers includes
- a process information receiving unit that receives information on the process from the client through the network;
- a determining unit that determines servers to execute the process from the servers based on the information on the process received by the process information receiving unit; and
- a server information transmitting unit that transmits information on the servers determined by the determining unit to the client, and
- the client includes
- a server information receiving unit that receives the information on the servers transmitted by the server information transmitting unit through the network; and
- a process request transmitting unit that transmits the information on the request for execution of the process to the servers received by the server information receiving unit.
2. The server/client system according to claim 1, wherein the determining unit includes
- a first distance calculating unit that calculates a first distance between an estimated point of a resource consumption obtained by adding a resource consumption of the process to a point representing the current resource consumption of each server, and a straight line connecting

the origin and the available maximum capacity of a parameter, in a space having parameters of resources in each server as axes; and

5 a second distance calculating unit that calculates a second distance between the estimated point of the resource consumption obtained by adding the resource consumption of the process to the point representing the current resource consumption of each server, and the origin, in the space having the parameters of the resources in each server as
10 axes, and determines servers to execute the process based on at least one of the first distance and the second distance.

3. The server/client system according to claim 2, wherein
15 the parameters include at least one of a load amount of a CPU, a load amount of a system memory, a load amount of a graphic processing unit, a load amount of a video memory, and a load amount of a network interface card.

20 4. A load distribution device used in a server/client system in which a plurality of servers and a plurality of clients are connected through a network, and the servers execute a process based on a process request from the client and transmit a process result to the client,
25 comprising

a process information receiving unit that receives information on the process from the client through the network;

a determining unit that determines servers to execute
30 the process from the servers based on the information on the process received by the process information receiving unit; and

a server information transmitting unit that transmits

information on the servers determined by the determining unit to the client.

5. The load distribution device according to claim 4,
5 wherein the determining unit includes

a first distance calculating unit that calculates a first distance between an estimated point of a resource consumption obtained by adding a resource consumption of the process to a point representing the current resource
10 consumption of each server, and a straight line connecting the origin and the available maximum capacity of a parameter, in a space having parameters of resources in each server as axes; and

a second distance calculating unit that calculates a
15 second distance between the estimated point of the resource consumption obtained by adding the resource consumption of the process to the point representing the current resource consumption of each server, and the origin, in the space having the parameters of the resources in each server as
20 axes, and determines servers to execute the process based on at least one of the first distance and the second distance.

6. The load distribution device according to claim 5,
25 wherein the parameters include at least one of a load amount of a CPU, a load amount of a system memory, a load amount of a graphic processing, a load amount of a video memory, and a load amount of a network interface card.

30 7. A load distribution method used in a server/client system in which a plurality of servers and a plurality of clients are connected through a network, and the servers execute a process based on a process request from the

client and transmit a process result to the client,
comprising

receiving information on the process from the client
through the network;

5 determining servers to execute the process from the
servers based on the information on the process received at
the receiving; and

transmitting the information on the request for
execution of the process to the servers determined at the
10 determining.

8. The load distribution method according to claim 7, the
determining includes

calculating a first distance between an estimated
15 point of a resource consumption obtained by adding a
resource consumption of the process to a point representing
the current resource consumption of each server, and a
straight line connecting the origin and the available
maximum capacity of a parameter, in a space having
20 parameters of resources in each server as axes; and

calculating a second distance between the estimated
point of the resource consumption obtained by adding the
resource consumption of the process to the point
representing the current resource consumption of each
25 server, and the origin, in the space having the parameters
of the resources in each server as axes, and determines
servers to execute the process based on at least one of the
first distance and the second distance.

30 9. The load distribution method according to claim 8,
wherein the parameters include at least one of a load
amount of a CPU, a load amount of a system memory, a load
amount of a graphic processing, a load amount of a video

memory, and a load amount of a network interface card.

10. A load distribution program for distributing loads of servers in a server/client system in which a plurality of
5 servers and a plurality of clients are connected through a network, and the servers execute a process based on a process request from the client and transmit a process result to the client, the load distribution program making the servers execute:
- 10 receiving information on the process from the client through the network;
- determining servers to execute the process from the servers based on the information on the process received at the receiving; and
- 15 transmitting the information on the request for execution of the process to the servers determined at the determining.
11. The load distribution program according to claim 10,
20 the determining includes
- calculating a first distance between an estimated point of a resource consumption obtained by adding a resource consumption of the process to a point representing the current resource consumption of each server, and a
25 straight line connecting the origin and the available maximum capacity of a parameter, in a space having parameters of resources in each server as axes; and
- calculating a second distance between the estimated point of the resource consumption obtained by adding the
30 resource consumption of the process to the point representing the current resource consumption of each server, and the origin, in the space having the parameters of the resources in each server as axes, and determines

servers to execute the process based on at least one of the first distance and the second distance.

12. The load distribution program according to claim 11,
5 wherein the parameters include at least one of a load amount of a CPU, a load amount of a system memory, a load amount of a graphic processing, a load amount of a video memory, and a load amount of a network interface card.